JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS 33, 482–496 (1971)

More on Linear Inequalities with Applications to Matrix Theory

ABRAHAM BERMAN*† AND ADI BEN-ISRAEL†

Department of Engineering Sciences, Northwestern University, Evanston, Illinois 60201

Submitted by Ky Fan

Revised December 2, 1969

Previous work [3, 4, 5] on solvability theorems for linear equations over cones and cones with interior is continued. Applications to matrix theory include a theorem of Bellman and Fan [2], and generalizations of Lyapunov theorem.

INTRODUCTION

This paper deals with linear equations over cones in finite dimensional spaces, and their applications to Hermitian matrices. It consists of four sections.

Section 0 collects the preliminaries on cones, polars and adjoints.

Section 1 is a survey of solvability theorems for the following systems:

(i)	$Ax = b,$	$x \in S$	(Theorem 1.1)
(ii)	$Ax = b,$	$x \in \operatorname{int} S$	(Theorem 1.3)
(iii)	$Ax \in \operatorname{int} S_1,$	$x \in \operatorname{int} S_2$	(Corollary 1.4)
(iv)	$Ax = 0,$	$0 \neq x \in S$	(Corollary 1.5)
(v)	$Ax = 0,$	$x \in \operatorname{int} S$	(Corollary 1.6)

where $A \in C^{m \times n}$, $b \in C^m$ and S, S_1, S_2 are suitable cones. In particular, the polyhedrality assumption common to previous theorems of alternative (e.g. [4] and the references therein) is relaxed in Corollaries 1.5 and 1.6.

Applications of these solvability theorems to matrix theory constitute the rest of the paper. The idea of using matrix cones to prove inertia and solva-

* His contribution is part of a Ph.D. Dissertation in Applied Mathematics, Northwestern University, 1970.

† Research partly supported by the National Science Foundation, projects GP 7550 and GP 13546 at Northwestern University.

bility theorems for matrices is not new, e.g. Schneider [13], Taussky [18] and Bellman and Fan [2]. That this is a very natural and useful idea is demonstrated by the ease with which matrix theorems are proved and extended here.

Section 2 uses Theorems 1.1 and 1.3 to prove the solvability theorem for Hermitian matrices of Bellman and Fan [2], which is a basis for a theory of linear inequalities and programming for Hermitian matrices. The proof here seems simpler than the original proof, and allows weaker assumptions.

Lyapunov theorem [1] and its various relatives and generalizations, e.g. [9, 11, 13, 14, 16, 17], are characterizations of solvability for the system (iii) where A is a suitable operator, S_1 is the cone of positive semidefinite matrices, and $S_2 = S_1$ or is the real space of Hermitian matrices. Theorems 3.1 and 3.2 in Section 3 are representative of this class of theorems, covered by the solvability theory of Section 1. Theorem 3.2 cannot be improved in the sense shown in Example 3.3.

0. Preliminaries

The notation of [3] will be adherred to in this paper.

0.1. Definitions. (a) A nonempty set S in C^n is:

(i) a *cone* if $0 \leqslant \lambda \Rightarrow \lambda S \subset S$

(ii) a *convex cone* if (i) and $S + S \subset S$

(iii) a *pointed cone* if (i) and $S \cap (-S) = \{0\}$.

(b) The *polar* S^* of a nonempty set $S \subset C^n$ is

$$S^* = \{y \in C^n : x \in S \Rightarrow \operatorname{Re}(y, x) \geqslant 0\}.$$

S^* is a closed convex cone and is the polar of the closed convex conical hull of S, i.e. of the smallest closed convex cone containing S. For other properties of polars see [3].

(c) The *interior* of S^*, int S^*, is given (algebraically) by:

$$\operatorname{int} S^* = \{y \in S^* : 0 \neq x \in S \Rightarrow \operatorname{Re}(y, x) > 0\}.$$

For a closed convex cone S, int S^* is nonempty iff S is pointed, [4].

$S = S^{**}$ iff S is a closed convex cone, [3] Theorem 1.5. For a closed convex cone S, the *interior* of S, int S, is therefore

$$\operatorname{int} S = \{x \in S : 0 \neq y \in S^* \Rightarrow \operatorname{Re}(y, x) > 0\}$$

and int $S \neq \phi$ iff S^* is pointed.

For polars of cartesian products the following lemma is easily proved.

0.2. LEMMA. *Let $S_j \subset C^{n(j)}$ be nonempty sets, $j = 1, \ldots, k$, and*

$$S = S_1 \times S_2 \times \cdots \times S_k \subset C^{\Sigma n(j)}.$$

Then

(a) $S^* = S_1^* \times S_2^* \times \cdots \times S_k^*$

(b) $\operatorname{int} S^* = \operatorname{int} S_1^* \times \operatorname{int} S_2^* \times \cdots \times \operatorname{int} S_k^*$. □

The reader may now proceed directly to Sec. 1. The applications to matrix theory given in Sections 2 and 3 require some additional preliminaries. For the inner product (,) in the above definitions, we use in $C^{n \times n}$

$$(X, Y) = \operatorname{tr} XY^H. \tag{1}$$

Let V denote the *real space of $n \times n$ Hermitian matrices*. The inner product (1) reduces on $V \times V$ to

$$(X, Y) = \operatorname{tr} XY \tag{2}$$

and is real valued.

0.3. EXAMPLES. We mention some cones in V, for later use.

(a) V is a (nonpointed) closed convex cone, and $V^* = \{0\}$.

(b) The set PSD of *positive semidefinite* matrices is a pointed closed convex cone. PSD is self-polar with respect to the inner product (2); i.e.

$$(PSD)^* = PSD, \qquad \text{e.g. ([5] Lemma 1).}$$

The set PD of *positive definite* matrices is the interior of PSD

$$PD = \operatorname{int} PSD, \qquad \text{e.g. ([5] Lemma 2).}$$

The cone PSD induces the following partial order in V:

$$X \geqslant Y \qquad \text{iff} \qquad X - Y \in PSD. \tag{3}$$

(c) The ray RI of nonnegative multiples of the identity matrix

$$RI = \{kI : k \geqslant 0\}$$

is a closed convex cone with empty interior. Its polar $(RI)^*$ is the (nonpointed) closed convex cone of all matrices in V with nonnegative trace

$$(RI)^* = \{X \in V : \operatorname{tr} X \geqslant 0\}.$$

0.4. DEFINITION. The *adjoint* T^* of a linear operator $T : C^{n \times n} \to C^{n \times n}$ is defined by

$$(TX, Y) = (X, T^*Y) \qquad \text{for all} \qquad X, Y \in C^{n \times n}. \tag{4}$$

It should be clear from the context whether a * denotes the polar of a set, or the adjoint of an operator.

The operators used below are of the form

$$TX = \sum_{j=1}^{k} A_j X B_j \qquad \text{where} \qquad A_j, B_j \in C^{n\times n}, \quad (j = 1,\dots, k). \tag{5}$$

The adjoint of this operator is given in terms of the conjugate transposes A_j^H, B_j^H of the matrices A_j, B_j $(j = 1,\dots, k)$:

0.5. LEMMA

$$T^*Y = \sum_{j=1}^{k} A_j^H Y B_j^H.$$

Proof. For any X, $Y \in C^{n\times n}$

$$\begin{aligned}(TX, Y) &= \operatorname{tr}\left[\sum_{j=1}^{k} A_j X B_j Y^H\right]\\ &= \operatorname{tr}\left[\sum_{j=1}^{k} X B_j Y^H A_j\right]\\ &= \left(X, \sum_{j=1}^{k} A_j^H Y B_j^H\right). \end{aligned}$$ □

1. LINEAR EQUATIONS OVER CONES

The basic solvability theorem for linear equations over cones is the following generalization of Farkas theorem:

1.1. THEOREM ([3] Theorem 2.4). *Let $A \in C^{m\times n}$, $b \in C^m$ and S a closed convex cone in C^n, and let $N(A) + S$ be closed. Then the following are equivalent:*

(a) *The system*

$$Ax = b, \qquad x \in S \tag{6}$$

is consistent.

(b) $A^H y \in S^* \Rightarrow \operatorname{Re}(b, y) \geqslant 0.$ □

The closedness of $N(A) + S$ is essential for Theorem 1.1 (for otherwise part (b) only characterizes asymptotic consistency, e.g. [3] Theorem 2.2).

A sufficient condition for $N(A) + S$ to be closed is that S be polyhedral, i.e. generated by finitely many vectors, ([3] Sec. 3). Another sufficient condition can be obtained from:

1.2. LEMMA. *Let L be a subspace and S a closed convex cone in C^n such that the intersection*

$$M = L \cap S$$

is a subspace of C^n. Then $L + S$ is closed.

Proof.[1]

(a) The lemma is proved first for the special case $M = \{0\}$.

Let $t = \lim_{k\to\infty}(l_k + s_k)$ where $\{l_k\} \subset L$, $\{s_k\} \subset S$. We have to show that $t \in L + S$. This is obvious if both sequences $\{l_k\}$, $\{s_k\}$ are bounded. The case of unbounded $\{l_k\}$ or $\{s_k\}$ is impossible for then the sequences

$$u_k = \frac{l_k}{\max\{\| l_k \|, \| s_k \|\}} \qquad \text{and} \qquad v_k = \frac{s_k}{\max\{\| l_k \|, \| s_k \|\}}$$

in L and S respectively, are bounded and contain convergent subsequences

$$u_{k_i} \to u \in L, \qquad v_{k_i} \to v \in S$$

where at least one of the vectors u, v is nonzero. Now

$$u + v = \lim_{k\to\infty} \frac{t}{\max\{\| u_k \|, \| v_k \|\}} = 0$$

$\therefore\ 0 \neq -u = v \in L \cap S = \{0\}$, a contradiction.

(b) The general case is reduced to the case (a). Let $L \cap S = M$.

(i) We show now that

$$S = M + S \cap M^{\perp} \tag{7}$$

$M + S \cap M^{\perp} \subset S$: obvious
$S \subset M + S \cap M^{\perp}$: Let $s \in S$. Then $s = x + y$, $x \in M$, $y \in M^{\perp}$.
Since $-x \in M \subset S$ it follows that

$$y = s - x \in S$$

$$\therefore \qquad s = x + y, \qquad x \in M, \qquad y \in S \cap M^{\perp}.$$

[1] This elementary proof, suggested by A. Charnes and A. Lent, is given for completeness. The lemma is a special case of [6], p. 78, exercise 10.

(ii) From (7) and $M \subset L$ it follows that

$$L + S = L + S \cap M^{\perp}.$$

Now $S \cap M^{\perp}$ is a closed convex cone, and

$$L \cap (S \cap M^{\perp}) = (L \cap S) \cap M^{\perp} = M \cap M^{\perp} = \{0\}.$$

Therefore $L + S$ is closed by part (a). □

The basic solvability theorem for linear equations with solutions in the interior of a cone is the following special case of Mazur's theorem, or the geometrical form of Hahn–Banach theorem e.g. ([6] p. 69) or ([12] p. 46).

1.3. THEOREM. *Let $A \in C^{m \times n}$, $b \in C^m$ and let S be a closed convex cone in C^n with nonempty interior. Then the following are equivalent.*

(a) *The system*

$$Ax = b, \qquad x \in \operatorname{int} S \tag{8}$$

is consistent

(b) $b \in R(A)$ *and* $0 \neq A^H y \in S^* \Rightarrow \operatorname{Re}(b, y) > 0$. □

A useful corollary of Theorem 1.3 is:

1.4. COROLLARY. *Let $T \in C^{m \times n}$, and let S_1 and S_2 be closed convex cones with nonempty interiors in C^m and C^n respectively. Then the following are equivalent:*

(a) *The system*

$$Tx \in \operatorname{int} S_1, \qquad x \in \operatorname{int} S_2 \tag{9}$$

is consistent.

(b) $-y \in S_1^*$, $T^H y \in S_2^* \Rightarrow y = 0$.

Proof. The system (9) is rewritten as:

$$(T, -I)\begin{pmatrix} x \\ z \end{pmatrix} = 0, \qquad \begin{pmatrix} x \\ z \end{pmatrix} \in \operatorname{int} S_2 \times \operatorname{int} S_1 = \operatorname{int}(S_2 \times S_1) \tag{10}$$

by Lemma 0.2.

By Theorem 1.3, the system (10) is consistent iff

$$0 \in R((T, -I)) \tag{11}$$

and

$$0 \neq \begin{pmatrix} T^H \\ -I \end{pmatrix} y \in (S_2 \times S_1)^* \Rightarrow \operatorname{Re}(0, y) > 0. \tag{12}$$

Now (11) is trivially satisfied, and the conclusion of the implication (12) is impossible. Therefore (10) is consistent iff

$$0 \neq \begin{pmatrix} T^H \\ -I \end{pmatrix} y \in (S_2 \times S_1)^*$$

is inconsistent. Using Lemma 0.2 this is equivalent to:

$$T^H y \in S_2^*, \qquad -y \in S_1^* \Rightarrow y = 0. \qquad \square$$

For $A \in C^{m \times n}$ and a closed convex cone $S \subset C^n$ it is natural to ask when does the system

$$Ax = 0, \qquad x \in S$$

possess nontrivial solutions.

Nontrivial here means nonzero or even that a solution lies in the (relative) interior of some face of S. The existence of nontrivial solutions is studied in theorems of alternative, each listing two systems such as

(I) $Ax = 0$, x nontrivial vector in S

(II) $A^H y$ nontrivial vector in S^*

exactly one of which is consistent. The theorems of alternative given in [4] and in the references therein are restricted to polyhedral cones.

The two theorems of alternative that follow are valid for general closed convex cones. This relaxation of the polyhedrality assumption is useful, e.g. in matrix applications where the nonpolyhedral cone of positive semidefinite matrices is used.

1.5. Corollary. *Let $A \in C^{m \times n}$ and S be a pointed closed convex cone in C^n. Then exactly one of the following systems is consistent.*

(I) $Ax = 0$, $0 \neq x \in S$

(II) $A^H y \in \operatorname{int} S^*$.

Proof. The system (II) can be rewritten as

(II′) $A^H y \in \operatorname{int} S^*$, $y \in \operatorname{int} C^m$.

The consistency of (II′) is equivalent, by Corollary 1.4, to

$$-x \in S, \qquad Ax \in (C^m)^* = \{0\} \Rightarrow x = 0$$

which completes the proof. $\square$

In the real case, for $S = R_+^n = S^*$ Corollary 1.5 reduces to the classical transposition theorem of Gordan [8].

A similar generalization of Stiemke's transposition theorem [15] is:

1.6. COROLLARY ([5], Theorem 3). *Let $A \in C^{m\times n}$ and S be a closed convex cone in C^n with nonempty interior. Then exactly one of the following systems is consistent:*

(I) $Ax = 0,\ x \in \operatorname{int} S$

(II) $0 \neq A^H y \in S^*$.

Proof. Follows from Theorem 1.3 by taking $b = 0$. □

2. A THEOREM OF BELLMAN AND FAN[2]

As a corollary of the solvability theorems of Sec. 1 we get the well-known theorem of Bellman and Fan ([2] Theorem 1) on linear inequalities in Hermitian matrices.

2.1. THEOREM (Bellman and Fan). *Assumptions*

(a) *Let $A_{ij} \in C^{n\times n}$ be arbitrary, let B_i, $C_j \in C^{n\times n}$ be Hermitian, ($i = 1,\ldots, p$; $j = 1,\ldots, q$) and let c be a real number.*

(b) *There exist positive definite Hermitian matrices Y_i ($i = 1,\ldots, p$) satisfying*

$$\sum_{i=1}^{p} (Y_i A_{ij} + A_{ij}^H Y_i) + C_j = 0, \qquad (j = 1,\ldots, q). \tag{13}$$

Conclusion

The following are equivalent.

(a) *The system*

$$\begin{gathered} \sum_{j=1}^{q} (A_{ij} X_j + X_j A_{ij}^H) \geqslant B_i, \qquad (i = 1,\ldots, p),^3 \\ \operatorname{tr} \sum_{j=1}^{q} C_j X_j \geqslant c \\ X_j = X_j^H, \qquad (j = 1,\ldots, q) \end{gathered} \tag{14}$$

is consistent.

(b) *For any m Hermitian matrices $D_i \geqslant O$ [3] and any number $d \geqslant 0$ the relations*

$$\sum_{i=1}^{p} (D_i A_{ij} + A_{ij}^H D_i) + dC_j = O, \qquad (j = 1,\ldots, q) \tag{15}$$

[2] This section benefitted from discussions with O. Taussky.

[3] Here $\geqslant$ is in the sense of (3), Sec. 0.

imply

$$\operatorname{tr} \sum_{i=1}^{p} D_i B_i + dc \leqslant 0. \tag{16}$$

Proof. For notational convenience, the proof is given in the case $p = q = 1$, where the indices of the matrices are omitted. The proof of the general case is similar.

For any $L \in C^{n \times n}$ the (Lyapunov) operator $T_L : V \to V$ is

$$T_L X = LX + XL^H. \tag{17}$$

The adjoint $T_L{}^*$ of T_L is by Lemma 0.5

$$\begin{aligned} T_L{}^* Y &= T_{L^H} Y \\ &= L^H Y + YL. \end{aligned} \tag{18}$$

For $p = q = 1$ the system (14) can be rewritten, by using (17) and the notations of Sec. 0, as:

$$T \begin{pmatrix} x \\ u \\ w \end{pmatrix} \equiv \begin{pmatrix} T_A & -I & O \\ T_C & O & -I \end{pmatrix} \begin{pmatrix} x \\ u \\ w \end{pmatrix} = \begin{pmatrix} B \\ \dfrac{2c}{n} I \end{pmatrix} \tag{19}$$

$$\begin{pmatrix} x \\ u \\ w \end{pmatrix} \in V \times PSD \times (RI)^* \equiv S.$$

If $N(T) + S$ is closed (for T and S defined by (19)) then the consistency of (19) is equivalent by Theorem 1.1 and examples 0.3(a), (b), (c) to:

$$\begin{pmatrix} T_A{}^* & T_C{}^* \\ -I & O \\ O & -I \end{pmatrix} \begin{pmatrix} -D \\ -\dfrac{d}{2} I \end{pmatrix} \in S^* = \{0\} \times PSD \times RI \tag{20}$$

implies

$$\operatorname{tr}(-D) B + (-d) c \geqslant 0. \tag{21}$$

For $p = q = 1$, (18) shows that (20) and (21) are equivalent to (15) and (16) respectively.

To complete the proof it therefore suffices to show that $N(T) + S$ is closed for which a sufficient conditon is, by Lemma 1.2, that $N(T) \cap S$ is a subspace. Indeed, the latter assertion follows from assumption (b), which for $p = q = 1$ states that the system

$$T_A{}^* Y = -C, \qquad Y \in \operatorname{int} PSD \tag{22}$$

is consistent, and by Theorem 1.3 is equivalent to

$$O \neq T_A X \in PSD \Rightarrow \operatorname{tr} CX < 0. \tag{23}$$

From (19)

$$N(T) = \left\{ \begin{pmatrix} X \\ T_A X \\ T_C X \end{pmatrix} : X \in V \right\}.$$

Thus $N(T) \cap S$ consists of the vectors $\begin{pmatrix} X \\ T_A X \\ T_C X \end{pmatrix}$ with $X \in V$,

$$\begin{aligned} &T_A X \in PSD, \qquad \text{and} \\ &T_C X \in (RI)^*. \end{aligned} \tag{24}$$

For such vectors $T_A X = O$ since otherwise $\operatorname{tr} CX < 0$ by (23), contradicting (24).

Therefore

$$N(T) \cap S = \left\{ \begin{pmatrix} X \\ O \\ T_C X \end{pmatrix} : X \in N(T_A) \right\}$$

is a subspace, completing the proof for the special case $p = q = 1$.

For the general case, the only modification needed in this proof is changing T and S to:

$$T = \begin{pmatrix} T_{A_{11}} & T_{A_{12}} & \cdots & T_{A_{1q}} & -I & & \\ \vdots & & & & & \ddots & \\ T_{A_{p1}} & T_{A_{p2}} & \cdots & T_{A_{pq}} & & -I & \\ T_{C_1} & T_{C_2} & \cdots & T_{C_q} & & & -I \end{pmatrix}$$

and

$$S = \underbrace{V \times \cdots \times V}_{q} \times \underbrace{PSD \times \cdots \times PSD}_{p} \times (RI)^*. \qquad \square$$

2.2. *Remark.* The above proof reveals that assumption (b) is used only to guarantee the closedness of $N(T) + S$, and therefore it can be replaced by

Assumption (b′): $N(T) + S$ is closed.

The following example shows that assumption (b′) is weaker than assumption (b).

EXAMPLE.

$$p = q = 1, \qquad A = I, \qquad C = O.$$

Assumption (b) calls for a positive definite $Y = O$, and therefore does no hold.

However assumption (b′) does hold, since

$$\begin{aligned} N(T) + S &= \left\{ \begin{pmatrix} X \\ 2X \\ O \end{pmatrix} : X \in V \right\} + V \times PSD \times (RI)^* \\ &= V \times V \times (RI)^* \end{aligned}$$

is closed.

3. Lyapunov Type Theorems

This section applies Corollary 1.4 to Hermitian matrices, by choosing S_1, S_2 and T, in Corollary 1.4, to be

$$S_1 = PSD, \qquad S_2 = PSD \quad \text{or} \quad V$$

and

$$T(X) = \sum_{i,j=1}^{s} d_{ij} A_i X A_j^H \tag{25}$$

where

$$(d_{ij}) \equiv D = D^H \tag{26}$$

and $A_1, A_2, \ldots, A_s$ are simultaneously triangulable, i.e. there is a nonsingular matrix Q such that

$$Q^{-1} A_i Q \equiv B_i = \begin{pmatrix} \lambda_1^{(i)} & & & \\ 0 & \lambda_2^{(i)} & & \\ \vdots & & \ddots & \\ 0 & \cdots & & \lambda_n^{(i)} \end{pmatrix}, \qquad (i = 1, \ldots, s). \tag{27}$$

This operator T was studied, using different methods, by Hill [9]. It includes as special cases the:

(i) *Lyapunov operator* $T(X) = AX + XA^H$, [1]

(ii) *Stein operator* $T(X) = X - CXC^H$, [14] and

(iii) *Schneider operator* $\quad T(X) = AXA^H - \sum_{i=1}^{p} C_i X C_i^H$

where A, $C_1, \ldots, C_p$ are simultaneously triangulable, [13].

The two choices of $S_2 = PSD$ and $S_2 = V$, result in [9] Theorems 4 and 3. Because their proofs, as consequences of Corollary 1.4, are essentially the same, both results are combined in the following theorem, where the case $S_2 = V$ is denoted by primed numbering of equations and by square brackets. The necessary condition in (b), believed to be new, is added for completeness.

3.1. THEOREM (Hill). *Let the operator* $T : V \to V$ *be defined by* (25), (26) *and* (27), *and let*

$$\varphi_k = \sum_{i,j=1}^{s} d_{ij} \lambda_k^{(i)} \overline{\lambda_k^{(j)}} \qquad (k = 1, \ldots, n). \tag{28}$$

Then:

(a) *A sufficient condition for the consistency of*

$$T(X) \in PD, \qquad X \in PD \tag{29}$$

$$[T(X) \in PD, \qquad X \in V] \tag{29'}$$

is

$$\varphi_k > 0, \qquad (k = 1, \ldots, n) \tag{30}$$

$$[\varphi_k \neq 0, \qquad (k = 1, \ldots, n)]. \tag{30'}$$

(b) *A necessary condition for the consistency of* (29) [(29′)] *is*

$$\varphi_n > 0 \tag{31}$$

$$[\varphi_n \neq 0]. \tag{31'}$$

(c) *If* $A_1, \ldots, A_s$ *are quasi commutative* (i.e. each A_i commutes with $A_j A_k - A_k A_j$, $(i, j, k = 1, \ldots, s)$) *then* (30) [(30′)] *is also a necessary condition for the consistency of* (29) [(29′)].

Proof. The consistency of (29) [(29′)] is equivalent, by Corollary 1.4, to

$$-T^*(Y) \in PSD, \qquad Y \in PSD \Rightarrow Y = O \tag{32}$$

$$[-T^*(Y) = O, \qquad Y \in PSD \Rightarrow Y = O] \tag{32'}$$

where by Lemma 0.5

$$T^*(Y) = \sum_{i,j=1}^{s} \overline{d_{ij}} A_i{}^H Y A_j$$

(a) We show that $\sim(32) \Rightarrow \sim(30)$ $[\sim(32') \Rightarrow \sim(30')]$ where $\sim$ denotes *negation.*

Let $O \neq Y \in PSD$ be such that $-T^*(Y) \in PSD[T^*(Y) = O]$. Then $Z = (z_{ij}) \equiv Q^H Y Q$ satisfies $O \neq Z \in PSD$ e.g. ([10], p. 84) and

$$G = (g_{ij}) \equiv Q^H T^*(Y) Q = \sum_{i,j=1}^{s} \overline{d_{ij}} B_i{}^H Z B_j$$

satisfies $-G \in PSD\,[G = O]$.

Let k be the first integer for which $z_{kk} > 0$. Then the first $(k-1)$ rows and columns of Z are zero (since $Z \in PSD$) and

$$g_{kk} = \sum_{i,j=1}^{s} \overline{d_{ij}} \overline{\lambda_k^{(i)}} \lambda_k^{(j)} z_{kk} = \varphi_k z_{kk}\,.$$

Therefore:

$$-G \in PSD \Rightarrow g_{kk} \leqslant 0 \Rightarrow \varphi_k \leqslant 0 \Rightarrow \sim(30)$$
$$[G = O \Rightarrow \varphi_k = 0 \Rightarrow \sim(30')].$$

(b) Let X be a solution of (29) [(29′)]. Then

$$W \equiv Q^{-1} X Q^{-H} \in PD[W \in V]$$

and

$$F \equiv Q^{-1} T(X) Q^{-H} = \sum_{i,j=1}^{s} d_{ij} B_i W B_j{}^H \in PD$$
$$\therefore \quad f_{nn} = \varphi_n w_{nn} > 0$$
$$\therefore \quad \varphi_n > 0 \qquad [\varphi_n \neq 0].$$

(c) Assuming quasi-commutativity we show

$$\sim(31) \Rightarrow \sim(32)\ [\sim(31') \Rightarrow \sim(32')].$$

Since $A_1, \ldots, A_s$ are quasi-commutative, so are $A_1{}^H, \ldots, A_s{}^H$. Thus for every $k = 1, \ldots, n$ there exists a common eigenvector u_k such that ([7])

$$A_i{}^H u_k = \lambda_k^{(i)} u_k \qquad (i = 1, \ldots, s).$$

Now

$$O \neq u_k u_k^H \in PSD \qquad (k = 1,..., n)$$

and

$$T^*(u_k u_k^H) = \sum_{i,j=1}^{s} \overline{d_{ij} \lambda_k^{(i)}} \lambda_k^{(j)} u_k u_k^H = \varphi_k u_k u_k^H.$$

Therefore for any $k = 1,..., n$:

$$\varphi_k \leqslant 0 \Rightarrow - T^*(u_k u_k^H) \in PSD \Rightarrow \sim (32)$$
$$[\varphi_k = 0 \Rightarrow T^*(u_k u_k^H) = 0 \Rightarrow \sim (32')]. \qquad \square$$

The relation of this theorem to the inertia theorems of Lyapunov [1], Stein [14], Schneider [13], Ostrowski and Schneider [11], and Taussky [16] is discussed in [9].

If the matrix D has exactly one positive eigenvalue, and $S_2 = PSD$ then more can be said about the operator T:

3.2. THEOREM. *Let T be as in Theorem* 3.1 *where D has exactly one positive eigenvalue. Let φ_k be defined by* (28). *Then the following are equivalent:*

(a) *The system*

$$T(X) \in PD, \qquad X \in PD$$

is consistent.

(b) *T is nonsingular and*

$$T(X) \in PD \Rightarrow X \in PD$$

(c) $\varphi_k > 0, \quad (k = 1,..., n)$.

Proof. Follows a theorem of Carlson ([9] p. 139) and a lemma of Schneider ([13] Lemma 1). $\square$

This theorem was proved by Schneider ([13] Theorem 1) for $D = (^1 \; {}_{-I})$ and by Taussky [17] for the Lyapunov operator. The equivalence of (a) and (c) was shown by Hill ([9] p. 140). Remark 3 in ([9] p. 141) follows from (b).

3.3. EXAMPLE. This example shows that Theorem 3.2 is not valid for matrices D with more than one positive eigenvalue, even if the matrices A_1 ,..., A_s are quasi-commutative:

$$D = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad A_1 = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}, \qquad A_2 = \begin{pmatrix} 1 & -a \\ 0 & 1 \end{pmatrix}.$$

Here A_1 and A_2 commute ($A_2 = A_1^{-1}$), but

$$T(X) = T\begin{pmatrix} x_1 & x_2 \\ \bar{x}_2 & x_3 \end{pmatrix} = A_1 X A_1{}^H + A_2 X A_2{}^H$$

$$= 2\begin{pmatrix} x_1 + |a|^2 x_3 & x_2 \\ \bar{x}_2 & x_3 \end{pmatrix}$$

satisfies (a) and (c), and does not satisfy (b) if $a \neq 0$.

References

1. R. Bellman, "Introduction to Matrix Analysis," McGraw-Hill, New York, 1960.
2. R. Bellman and K. Fan, On systems of linear inequalities in Hermitian matrix variables, *in* Proceedings of Symposia in Pure Mathematics, *Volume VII: Convexity*, (V. Klee, Ed.), pp. 1–11, American Math. Soc., Providence, R. I., 1963.
3. A. Ben-Israel, Linear equations and inequalities on finite dimensional, real or complex, vector spaces: A unified theory, *J. Math. Anal. Appl.* **27** (1969), 367–389.
4. A. Ben-Israel, Theorems of the alternative for complex linear inequalities, *Israel J. Math.* **7** (1969), 129–136.
5. A. Berman and A. Ben-Israel, Linear equations over cones with interior: A solvability theorem with applications to matrix theory, Report No. 69-1, Series in Applied Mathematics, The Technological Institute, Northwestern University, Evanston, Ill., 1969.
6. N. Bourbaki, Espaces vectoriels topologiques, Hermann & Cie, Paris, 1953.
7. M. P. Drazin, J. W. Dungey, and K. W. Gruenberg, Some theorems on commutative matrices, *J. London Math. Soc.* **26** (1951), 221–228.
8. P. Gordan, Über die Auflösungen linearer Gleichungen mit reelen Coefficienten, *Math. Ann.* **6** (1873), 23–28.
9. R. D. Hill, Inertia theorems for simultaneously triangulable complex matrices, *J. Lin. Algebra and Its Applications* **2** (1969), 131–142.
10. M. Marcus and M. Minc, A survey of matrix theory and matrix inequalities, Allyn and Bacon, Boston, Mass., 1964.
11. A. Ostrowski and H. Schneider, Some theorems on the inertia of general matrices, *J. Math. Anal. Appl.* **4** (1962), 72–84.
12. H. H. Schaefer, Topological vector spaces, Macmillan, New York, 1966.
13. H. Schneider, Positive operators and an inertia theorem, *Numer. Math.* **7** (1965), 11–17.
14. P. Stein, Some general theorems on iterants, *J. Research, Nat. Bur. Standards* **48** (1952), 82–83.
15. E. Stiemke, Über positive Lösungen homogener linearer Gleichungen, *Math. Ann.* **76** (1915), 340–342.
16. O. Taussky, A generalization of a theorem of Lyapunov, *J. Soc. Ind. Appl. Math.* **9** (1961), 640–643.
17. O. Taussky, Matrices C with $C^n \to 0$, *J. Algebra* **1** (1964), 5–10.
18. O. Taussky, Positive definite matrices and their role in the study of the characteristic roots of general matrices, *Advan. Math.* **2** (1968), 175–186.